

Commodity Machine Learning

Andreas Müller
Amazon, scikit-learn

✉ importamueller@gmail.com

🐦 [@t3kcit](#)

🐙 [@amueller](#)

📄 peekaboo-vision.blogspot.com

We Want You!



To Apply Machine Learning!

What ML can do for you

Classification

Hi Andy,

I just received an email from the first tutorial speaker, presenting right before you, saying he's ill and won't be able to make it.

I know you have already committed yourself to two presentations, but is there anyway you could increase your tutorial time slot, maybe just offer time to try out what you've taught? Otherwise I have to do some kind of modern dance interpretation of Python in data :-)
-Leah

Hi Andreas,

I am very interested in your Machine Learning background. I work for X Recruiting who have been engaged by Z, a worldwide leading supplier of Y. We are expanding the core engineering team and we are looking for really passionate engineers who want to create their own story and help millions of people.

Can we find a time for a call to chat for a few minutes about this?

Thanks

Classification

Hi Andy,

I just received an email from the first tutorial speaker, presenting himself, saying he's ill and won't be

I know you have a lot to offer yourself to two presentations. If you could increase your audience, maybe just offer time to try out what you've taught? Otherwise I have to do some kind of modern dance interpretation of Python in data :-)
-Leah



Hi Andreas,

I am very interested in your Machine Learning background. I work for X Recruiting who have been engaged by Z, a worldwide leading supplier of Y. We are expanding the core engineering team and we are looking for really passionate engineers who can tell their own story and help millions of

Can we find a time for a few minutes about this?

Thanks



Classification

SPONSORED 

[Create Ad](#)

Online Essen bestellen

lieferheld.de



Neu mit Lieferheld: PLZ eingeben, Liefersdienst finden und genießen!

Globaler Chauffeurservice

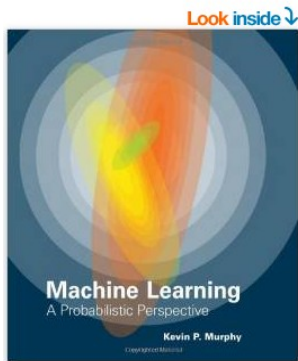
blacklane.com



Fahren Sie eine Klasse besser zu günstigen Preisen – Blacklane ist weltweit verfügbar!

97,974 people like this

Recommendations



Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) Hardcover

by Kevin P. Murphy (Author)

★★★★★ 31 customer reviews

See all 2 formats and editions

Kindle
\$64.06

Read with our free app

Hardcover
\$77.41

20 Used from \$67.99

42 New from \$69.66

Today's Web-enabled deluge of electronic data calls for automated methods of data analysis. Machine learning provides these, developing methods that can automatically detect patterns in data and then use the uncovered patterns to predict future data. This textbook offers a comprehensive and self-contained introduction to the field of machine learning, based on a unified, probabilistic approach. The coverage combines breadth and depth, offering necessary background material on such topics as probability, optimization, and linear algebra as well as discussion of recent developments in the field, including conditional random fields, L1 regularization, and deep learning. The book is written in an informal, accessible style, complete with pseudo-code for the most important algorithms. All topics are copiously illustrated with color images and worked examples drawn from such application domains as

Read more



See all 1 image(s)

Share

Buy New \$77.41

Qty: 1

List Price: \$99.99

Save: \$12.59 (14%)

FREE Shipping

In Stock.

Ships from and sold by Amazon.com. Gift-wrap available.



Add to Cart

Sign in to turn on 1-click ordering

Want it Wednesday, June 11? Order within 11 hrs 27 mins and choose AmazonGlobal Priority Shipping at checkout. Details

Buy Used \$71.98

Add to Wish List

Trade in your item
Get a \$35.10
Gift Card.

Trade In

Learn More

Have one to sell?

Sell on Amazon

Frequently Bought Together



Price for all three: \$249.90

Add all three to Cart

Add all three to Wish List

Show availability and shipping details

✓ **This Item:** Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) by Kevin P. Murphy Hardcover \$77.41

✓ Pattern Recognition and Machine Learning (Information Science and Statistics) by Christopher M. Bishop Hardcover \$85.46

✓ Probabilistic Graphical Models: Principles and Techniques (Adaptive Computation and Machine Learning ... by Daphne Koller Hardcover \$87.03

Customers Who Bought This Item Also Bought

Page 1 of 12



Learning From Data
Yaser S. Abu-Mostafa



Pattern Recognition and
Machine Learning ...



Probabilistic Graphical
Models: Principles ...



The Elements of Statistical
Learning: ...



Foundations of Machine
Learning (Adaptive ...



Machine Learning: The Art
and Science of ...



Building Machine Learning
Systems with ...

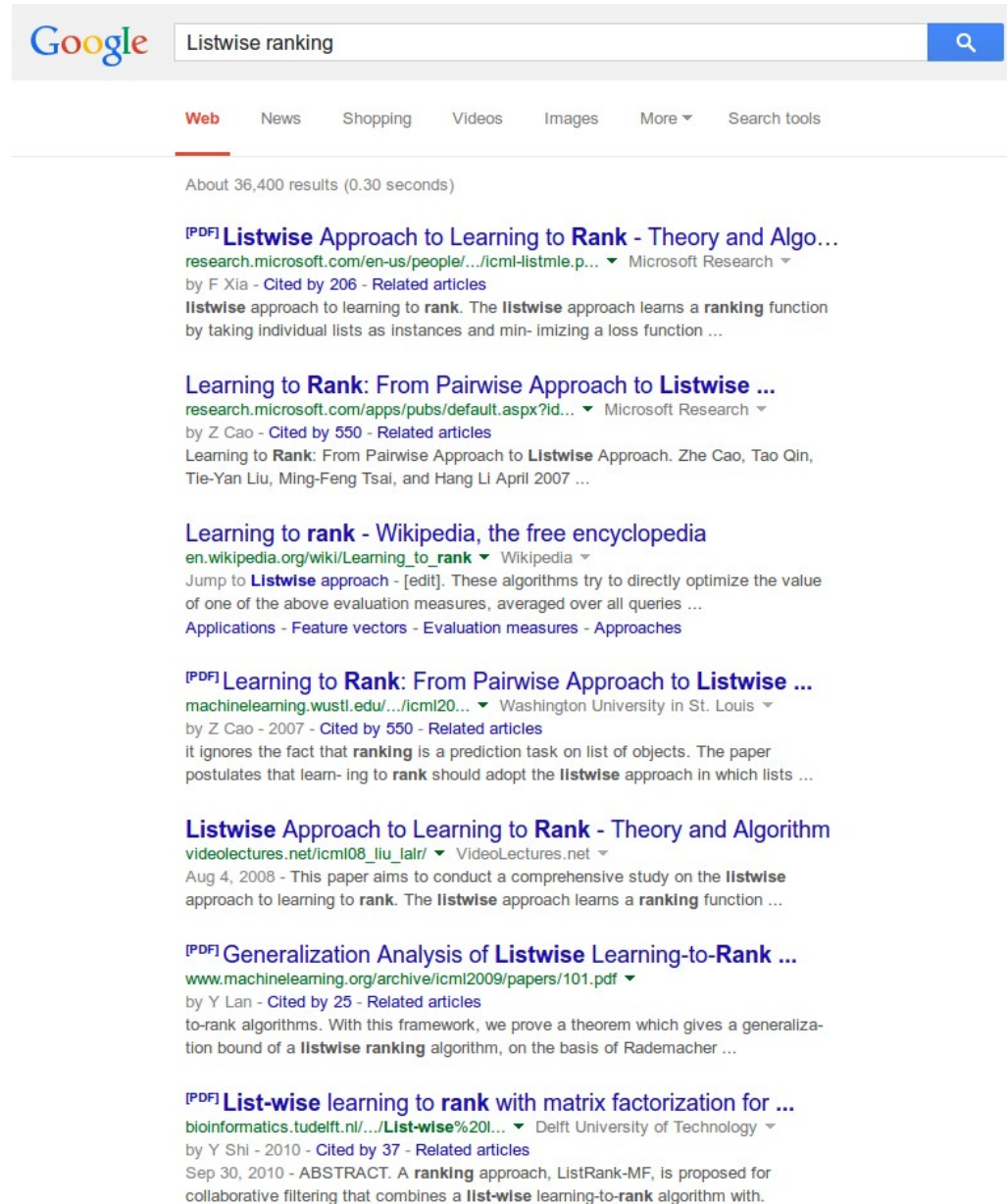


Understanding Machine
Learning: From Theory to ...



An Introduction to
Statistical Learning: ...

Ranking



The image shows a Google search results page for the query "Listwise ranking". The search bar at the top contains the text "Listwise ranking" and a magnifying glass icon. Below the search bar, there are tabs for "Web", "News", "Shopping", "Videos", "Images", "More", and "Search tools". The "Web" tab is selected. The search results are displayed below the tabs, showing about 36,400 results in 0.30 seconds. The first result is a PDF titled "Listwise Approach to Learning to Rank - Theory and Algo..." from Microsoft Research, by F Xia, cited by 206. The second result is "Learning to Rank: From Pairwise Approach to Listwise ..." from Microsoft Research, by Z Cao, cited by 550. The third result is "Learning to rank - Wikipedia, the free encyclopedia", which includes a jump to the "Listwise approach" section. The fourth result is "Learning to Rank: From Pairwise Approach to Listwise ..." from Washington University in St. Louis, by Z Cao, cited by 550. The fifth result is "Listwise Approach to Learning to Rank - Theory and Algorithm" from VideoLectures.net, dated Aug 4, 2008. The sixth result is a PDF titled "Generalization Analysis of Listwise Learning-to-Rank ..." from machinelearning.org, by Y Lan, cited by 25. The seventh result is "List-wise learning to rank with matrix factorization for ..." from bioinformatics.tudelft.nl, by Y Shi, cited by 37.

Google

Listwise ranking

Web News Shopping Videos Images More Search tools

About 36,400 results (0.30 seconds)

Listwise Approach to Learning to Rank - Theory and Algo...
[research.microsoft.com/en-us/people/.../icml-listmle.p...](#) Microsoft Research
by F Xia - Cited by 206 - Related articles
Listwise approach to learning to rank. The Listwise approach learns a ranking function by taking individual lists as instances and minimizing a loss function ...

Learning to Rank: From Pairwise Approach to Listwise ...
[research.microsoft.com/apps/pubs/default.aspx?id...](#) Microsoft Research
by Z Cao - Cited by 550 - Related articles
Learning to Rank: From Pairwise Approach to Listwise Approach. Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li April 2007 ...

Learning to rank - Wikipedia, the free encyclopedia
[en.wikipedia.org/wiki/Learning_to_rank](#) Wikipedia
Jump to Listwise approach - [edit]. These algorithms try to directly optimize the value of one of the above evaluation measures, averaged over all queries ...
Applications - Feature vectors - Evaluation measures - Approaches

Learning to Rank: From Pairwise Approach to Listwise ...
[machinelearning.wustl.edu/.../icml20...](#) Washington University in St. Louis
by Z Cao - 2007 - Cited by 550 - Related articles
It ignores the fact that ranking is a prediction task on list of objects. The paper postulates that learning to rank should adopt the Listwise approach in which lists ...

Listwise Approach to Learning to Rank - Theory and Algorithm
[videlectures.net/icml08_liu_lai/](#) VideoLectures.net
Aug 4, 2008 - This paper aims to conduct a comprehensive study on the Listwise approach to learning to rank. The Listwise approach learns a ranking function ...

Generalization Analysis of Listwise Learning-to-Rank ...
[www.machinelearning.org/archive/icml2009/papers/101.pdf](#)
by Y Lan - Cited by 25 - Related articles
to-rank algorithms. With this framework, we prove a theorem which gives a generalization bound of a Listwise ranking algorithm, on the basis of Rademacher ...

List-wise learning to rank with matrix factorization for ...
[bioinformatics.tudelft.nl/.../List-wise%20l...](#) Delft University of Technology
by Y Shi - 2010 - Cited by 37 - Related articles
Sep 30, 2010 - ABSTRACT. A ranking approach, ListRank-MF, is proposed for collaborative filtering that combines a List-wise learning-to-rank algorithm with.

Applying machine learning is easy.

Applying machine learning is easy.
But it should be easier!

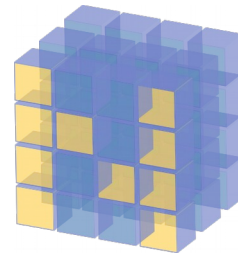


python

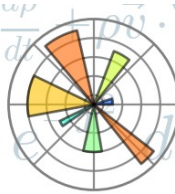
IP[y]: IPython
Interactive Computing



SciPy



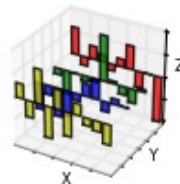
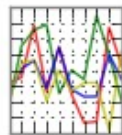
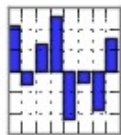
NumPy



matplotlib

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



IP[y]: IPython Interactive Computing

```
Chapter 4 - Working With Text Data. - Mozilla Firefox
File Edit View History Bookmarks Tools Help
IPy Chapter 4 - Workin...
127.0.0.1:8888/030923ca-af46-4fc6-86fc-1f0b18130e27#
DuckDuckGo
IP[y]: Notebook Chapter 4 - Working With Text Data. (autosaved)
File Edit View Insert Cell Kernel Help
Code Cell Toolbar: None

In [22]: from sklearn.svm import LinearSVC
svm = LinearSVC(C=0.01)

In [23]: svm.fit(X_train, y_train)

Out[23]: LinearSVC(C=0.01, class_weight=None, dual=True, fit_intercept=True,
intercept_scaling=1, loss='l2', multi_class='ovr', penalty='l2',
random_state=None, tol=0.0001, verbose=0)

In [24]: svm.score(X_train, y_train)

Out[24]: 0.88421586014694709

In [25]: svm.score(X_test, y_test)

Out[25]: 0.83679637325273892

In [26]: y_test_pred = svm.predict(X_test)

In [27]: from sklearn.metrics import classification_report

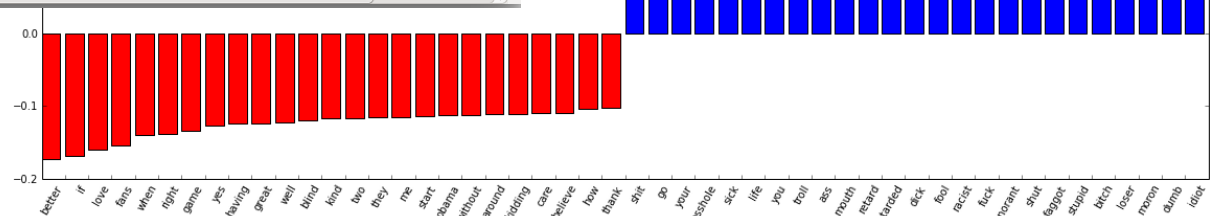
In [28]: print(classification_report(y_test, y_test_pred))

              precision    recall  f1-score   support

     0       0.84         0.96         0.90       1954
     1       0.82         0.48         0.61         693

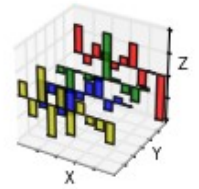
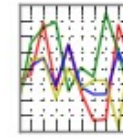
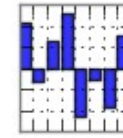
 avg / total       0.83         0.84         0.82       2647

In [29]: coef = svm.coef_.ravel()
positive_coefficients = np.argsort(coef)[-25:]
negative_coefficients = np.argsort(coef)[:25]
interesting_coefficients = np.hstack([negative_coefficients, positive_coefficients])
```



pandas

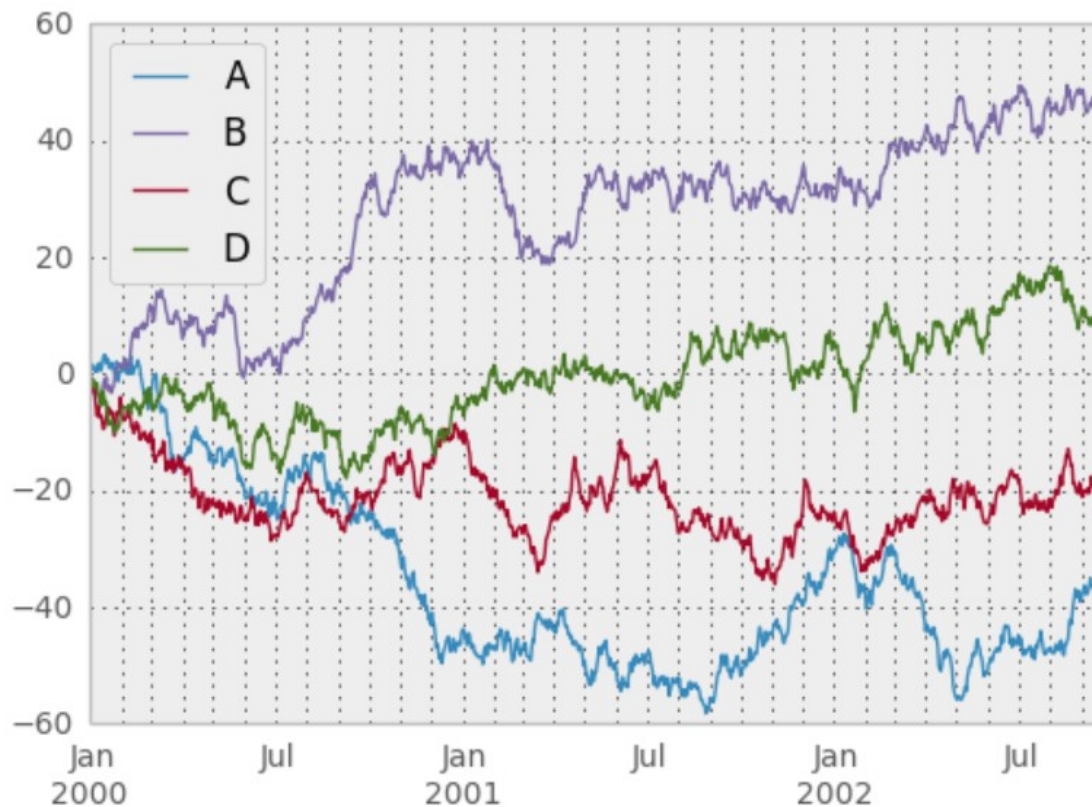
$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$

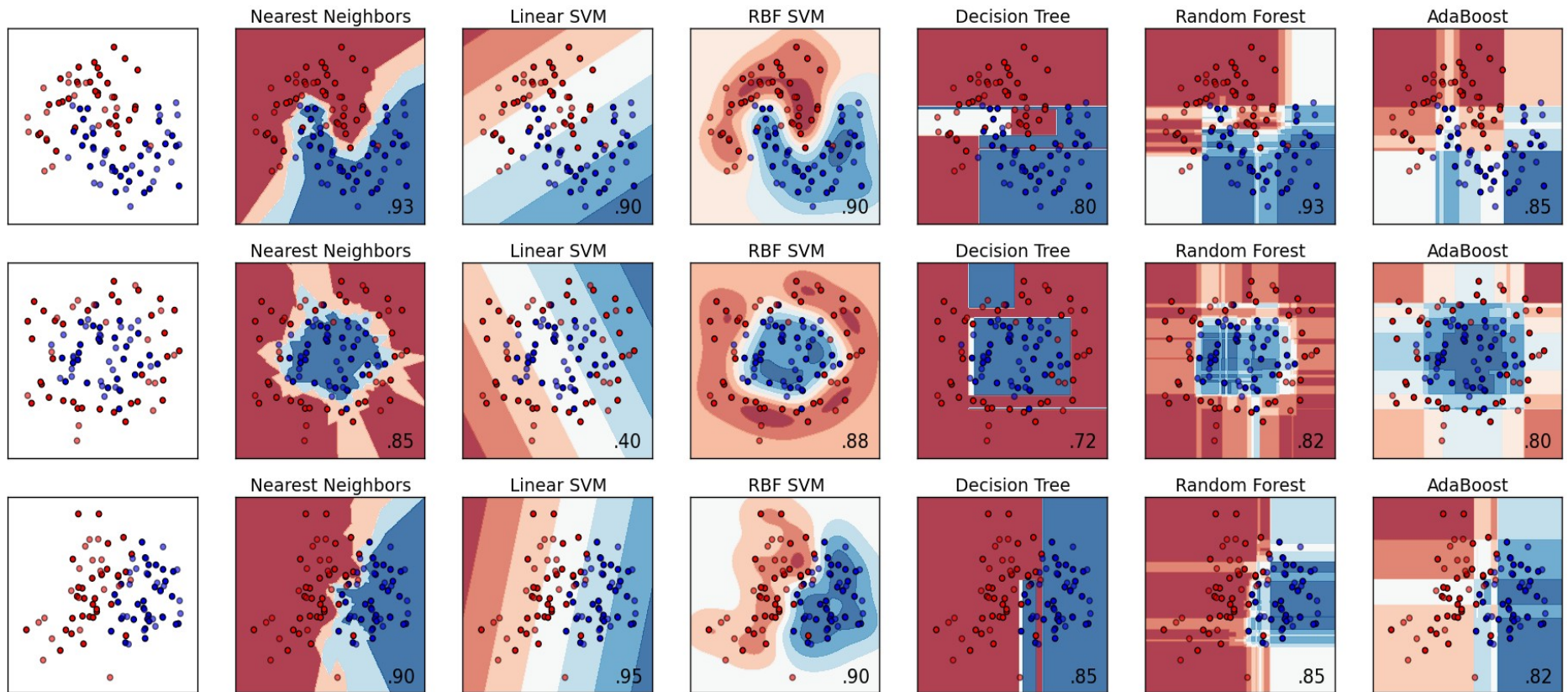


```
In [6]: df = DataFrame(randn(1000, 4), index=ts.index, columns=list('ABCD'))
```

```
In [7]: df = df.cumsum()
```

```
In [8]: plt.figure(); df.plot();
```





lovely



500+ research papers

```
from sklearn.ensemble import RandomForestClassifier
clf = RandomForestClassifier()
clf.fit(X_train, y_train)
clf.predict(X_test)
```

Fully Functional Spam Classifier

```
from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer
from pipeline import make_pipeline

spam_classifier = make_pipeline(CountVectorizer(),
                                MultinomialNB())
spam_classifier.fit(email_texts, is_spam)

spam_classifier.predict(new_emails)
```

Generalized Linear Models

Support Vector Machines

Stochastic Gradient Descent

Nearest Neighbors

Gaussian Processes

CCA

Naive Bayes

Decision Trees

Ensemble methods

Multiclass and multilabel algorithms

Clustering

Matrix Factorization

Manifold Learning

Mixture Models

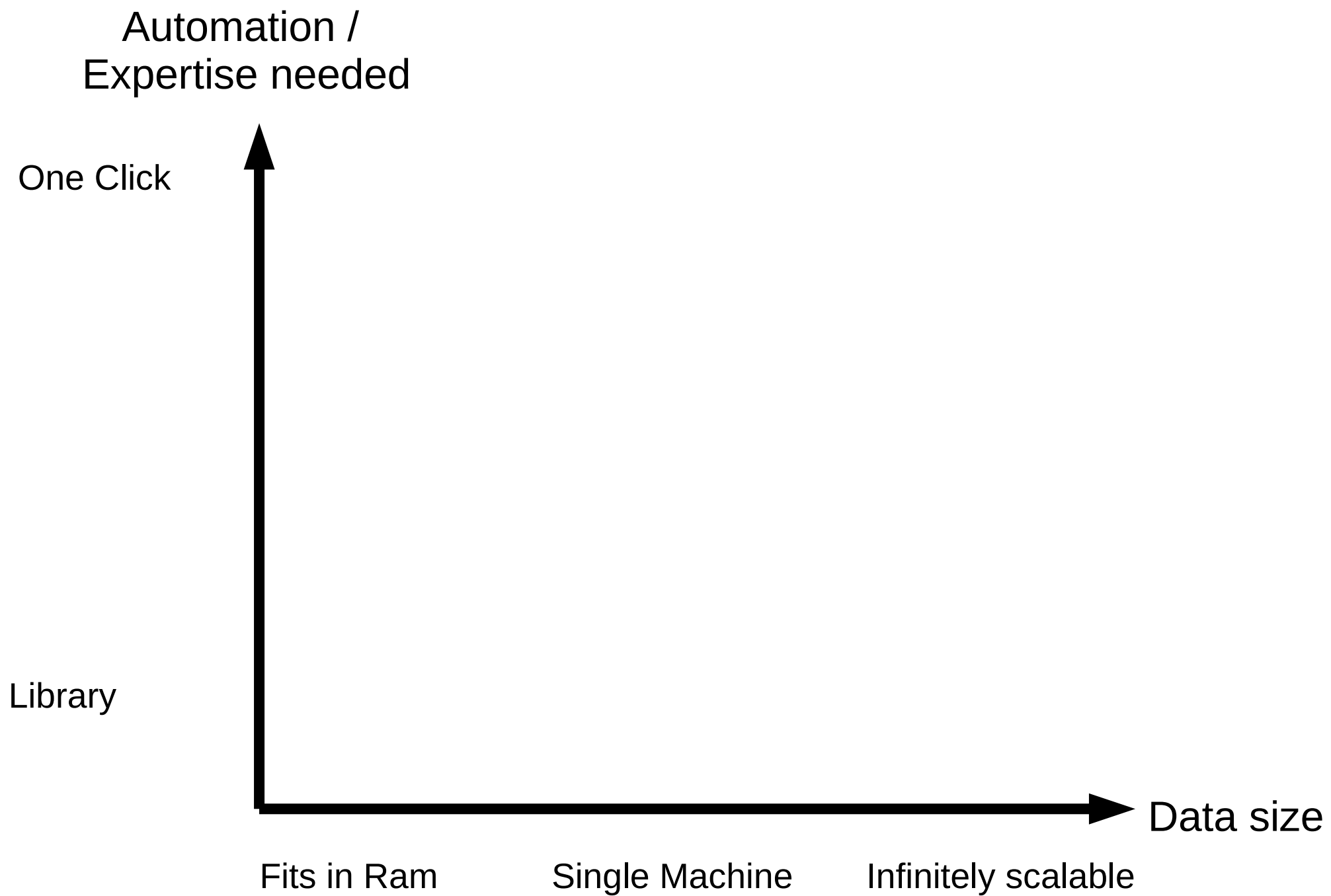
“The scikit-learn tutorials / documentation is so good, one doesn't need a textbook anymore to learn a new machine learning method.”

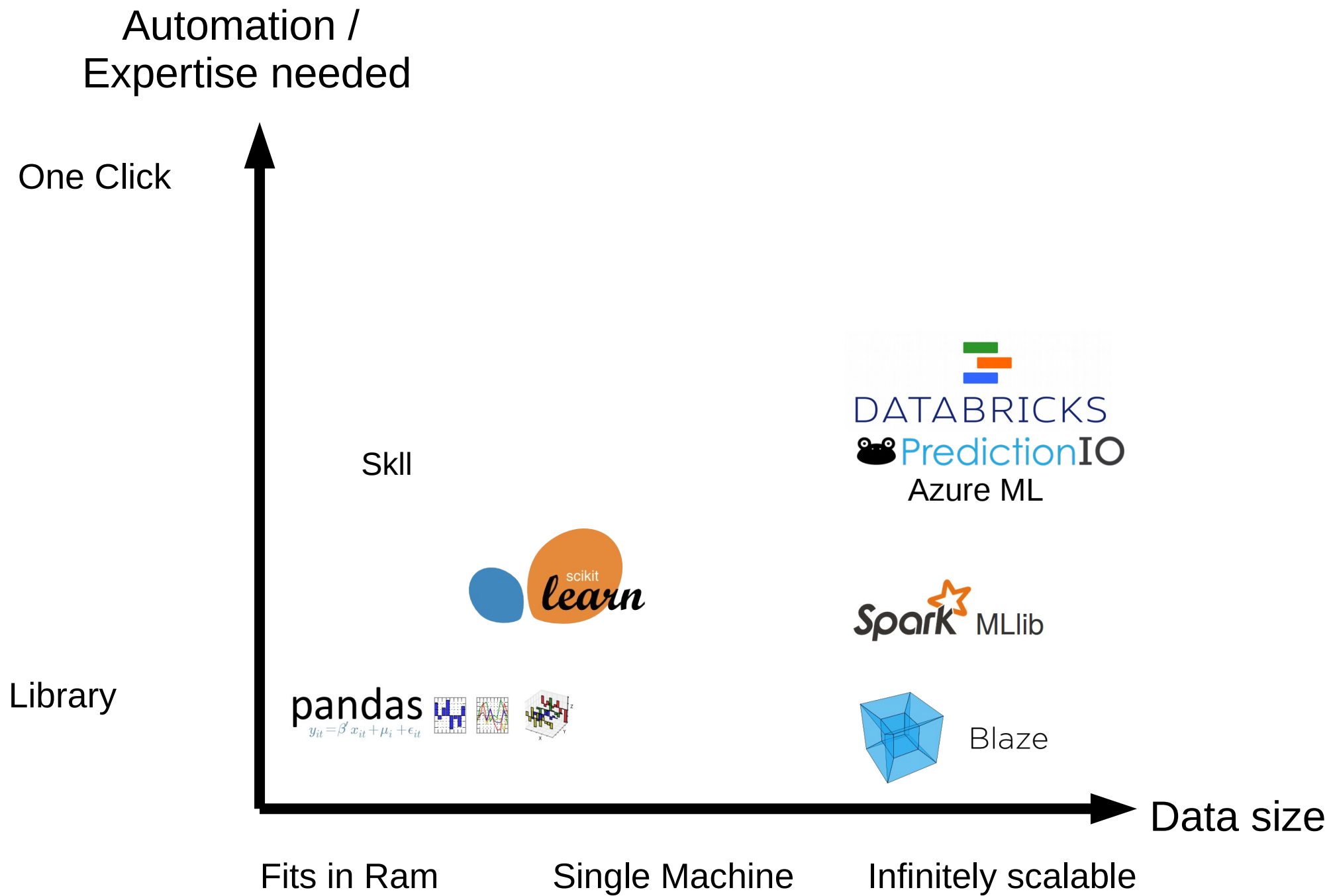
This is not enough!

Automation /
Expertise needed



Data size





Why a single machine is (usually) enough

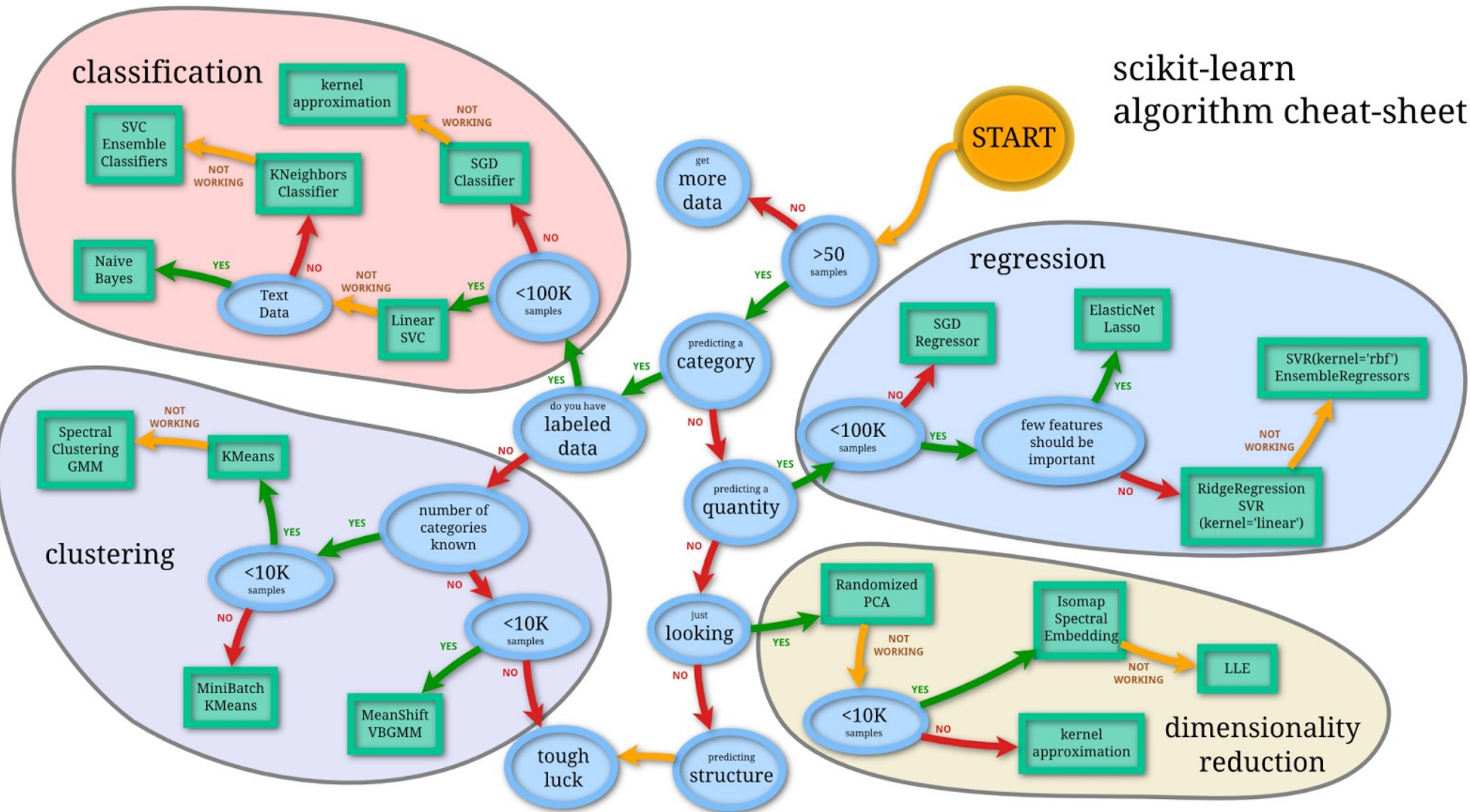
```
1 [|||||||||||||||||||||||||||||||||||||||||100.0%]
2 [|||||||||||||||||||||||||||||||||||||||||100.0%]
3 [|||||||||||||||||||||||||||||||||||||||||100.0%]
4 [|||||||||||||||||||||||||||||||||||||||||100.0%]
5 [|||||||||||||||||||||||||||||||||||||||||100.0%]
6 [|||||||||||||||||||||||||||||||||||||||||100.0%]
7 [|||||||||||||||||||||||||||||||||||||||||100.0%]
8 [|||||||||||||||||||||||||||||||||||||||||100.0%]
9 [|||||||||||||||||||||||||||||||||||||||||100.0%]
10 [|||||||||||||||||||||||||||||||||||||||||100.0%]
11 [|||||||||||||||||||||||||||||||||||||||||100.0%]
12 [|||||||||||||||||||||||||||||||||||||||||100.0%]
13 [|||||||||||||||||||||||||||||||||||||||||100.0%]
14 [|||||||||||||||||||||||||||||||||||||||||100.0%]
15 [|||||||||||||||||||||||||||||||||||||||||100.0%]
16 [|||||||||||||||||||||||||||||||||||||||||100.0%]
17 [|||||||||||||||||||||||||||||||||||||||||100.0%]
18 [|||||||||||||||||||||||||||||||||||||||||100.0%]
19 [|||||||||||||||||||||||||||||||||||||||||100.0%]
20 [|||||||||||||||||||||||||||||||||||||||||100.0%]
21 [|||||||||||||||||||||||||||||||||||||||||100.0%]
22 [|||||||||||||||||||||||||||||||||||||||||100.0%]
23 [|||||||||||||||||||||||||||||||||||||||||100.0%]
24 [|||||||||||||||||||||||||||||||||||||||||100.0%]
25 [|||||||||||||||||||||||||||||||||||||||||100.0%]
26 [|||||||||||||||||||||||||||||||||||||||||100.0%]
27 [|||||||||||||||||||||||||||||||||||||||||100.0%]
28 [|||||||||||||||||||||||||||||||||||||||||100.0%]
29 [||| |||||||||||||||||||||||||||||||||||||||100.0%]
30 [|||||||||||||||||||||||||||||||||||||||||100.0%]
31 [|||||||||||||||||||||||||||||||||||||||||100.0%]
32 [|||||||||||||||||||||||||||||||||||||||||100.0%]
Mem[||||||| |23164/245759M]
Swp[0/0MB]
```

Smart, not Big

Why we need open box methods

Why we need black-box methods

scikit-learn algorithm cheat-sheet



MLService

predict

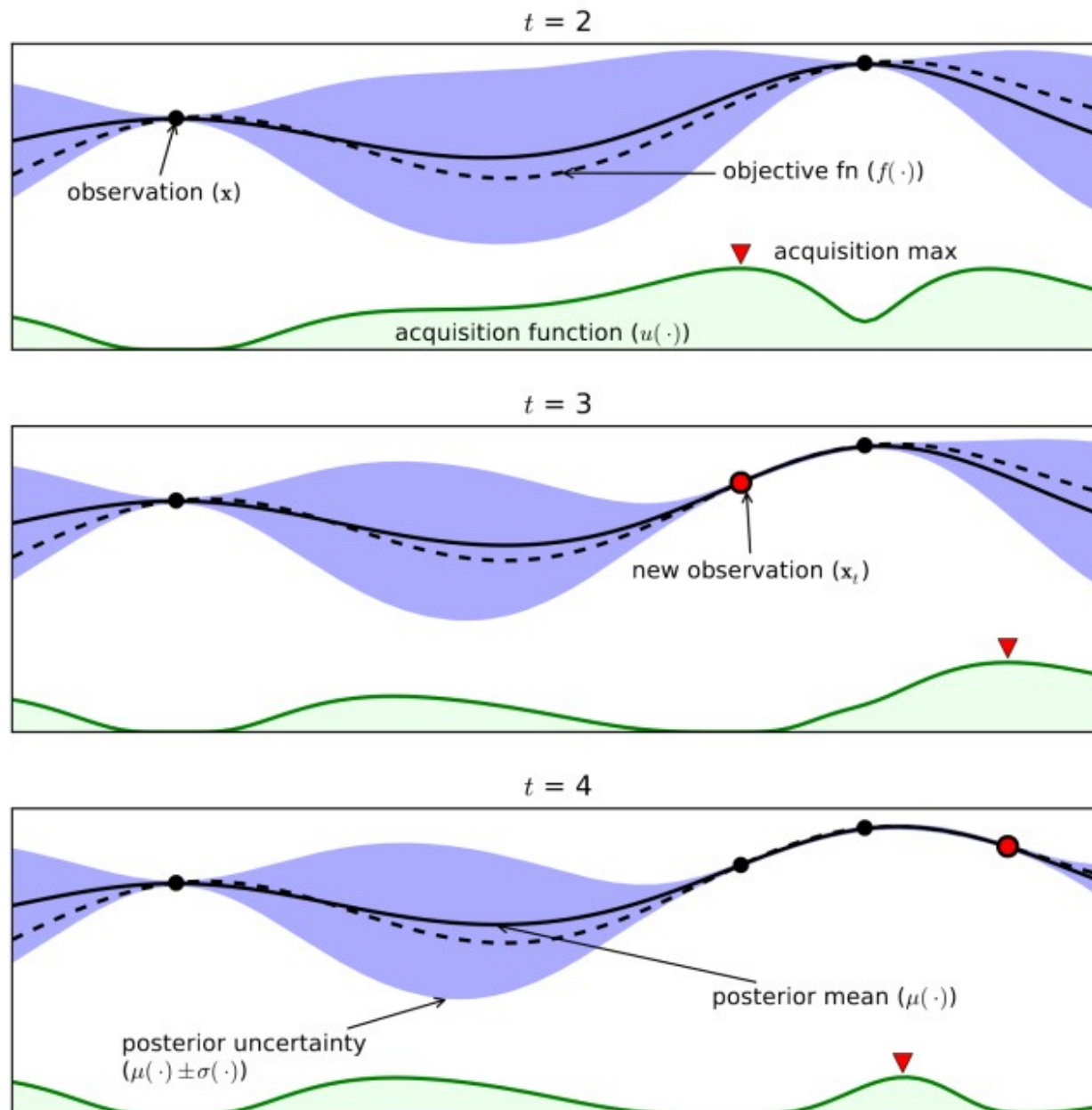
Hyperparameter Optimization

Spearmint

Hyperopt

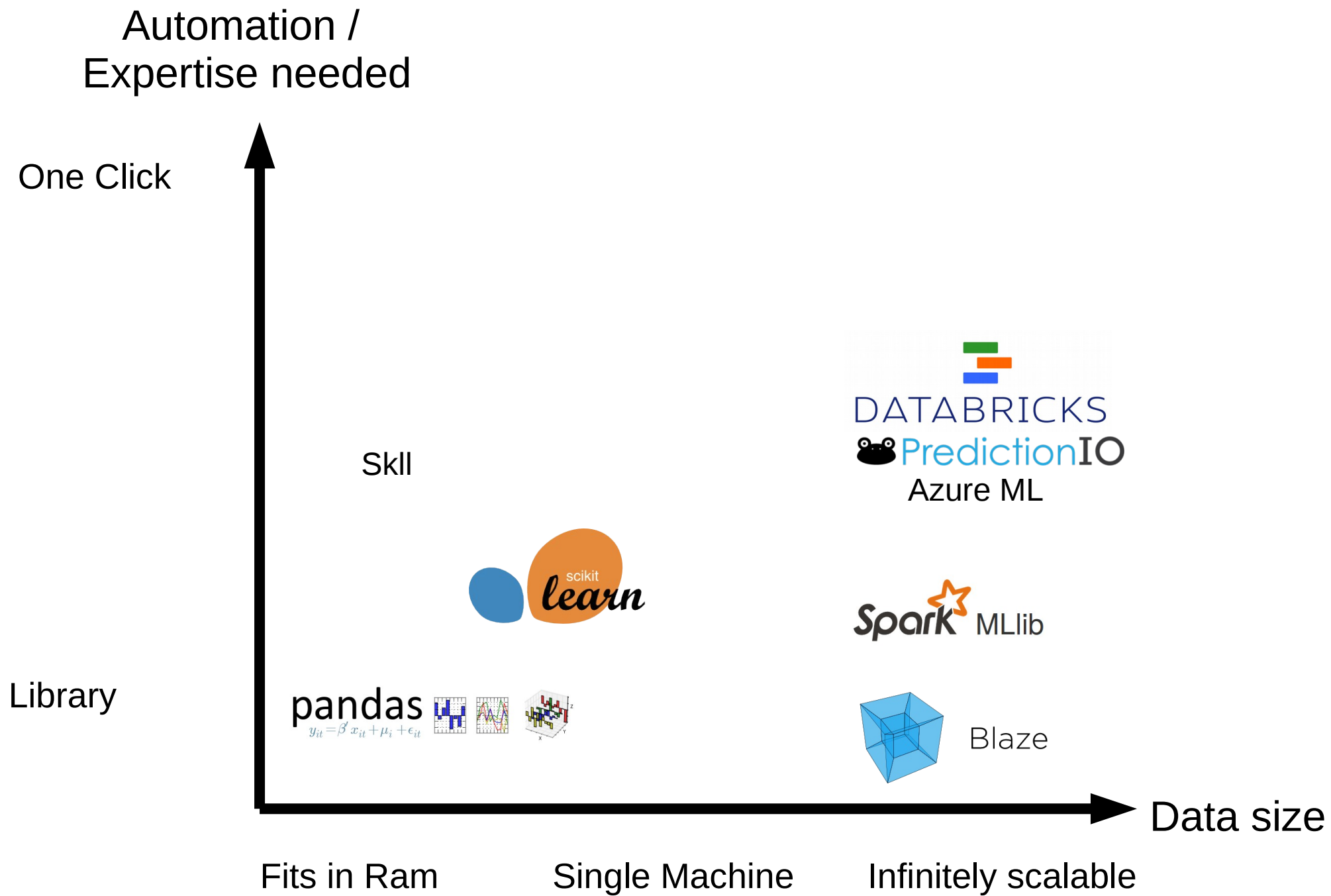
smac

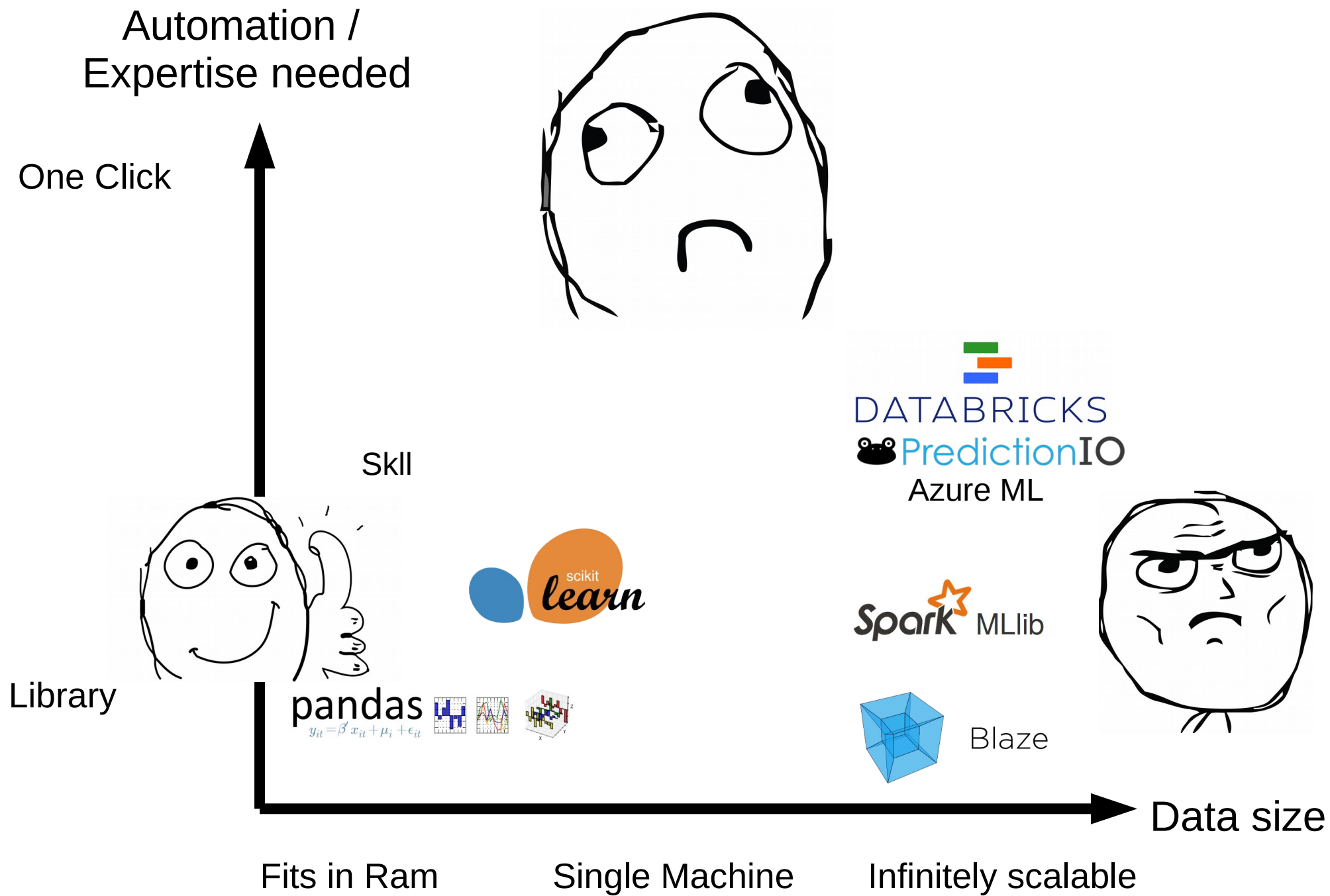
Bayesian Optimization



From Eric Brochu, Vlad M. Cora and Nando de Freitas

Why we need to scale beyond a single machine





Thank you.

Andreas Müller

✉ importamueller@gmail.com

🐦 [@t3kcit](#)

🐙 [@amueller](#)

📄 peekaboo-vision.blogspot.com